Validation of non-stationary precipitation series for site-specific impact assessment: comparison of two statistical downscaling techniques

Donal Mullan[1], Jie Chen[2], Xunchang John Zhang[3]

[1] School of Geography, Archaeology and Palaeoecology, Queen's University Belfast, Elmwood Avenue, Belfast BT7 1NN, Co. Antrim, Northern Ireland

[2] Department of Construction Engineering, École de Technologie Supérieure, Université du Québec, Montreal, Canada

[3] USDA-ARS, Grazinglands Research Laboratory, El Reno, Oklahoma, USA

Abstract

Statistical downscaling (SD) methods have become a popular, low-cost and accessible means of bridging the gap between the coarse spatial resolution at which climate models output climate scenarios and the finer spatial scale at which impact modellers require these scenarios, with various different SD techniques used for a wide range of applications across the world. This paper compares the Generator for Point Climate Change (GPCC) model and the Statistical DownScaling Model (SDSM) – two contrasting SD methods – in terms of their ability to generate precipitation series under non-stationary conditions across ten contrasting global climates. The mean, maximum and a selection of distribution statistics as well as the cumulative frequencies of dry and wet spells for four different temporal resolutions were compared between the models and the observed series for a validation period. Results indicate that both methods can generate daily precipitation series that generally closely mirror observed series for a wide range of non-stationary climates. However, GPCC tends to overestimate higher precipitation amounts, whilst SDSM tends to underestimate these. This infers that GPCC is more likely to overestimate the effects of precipitation on a given impact sector, whilst SDSM is likely to underestimate the effects. GPCC performs better than SDSM in reproducing wet and dry day frequency, which is a key advantage for many impact sectors. Overall, the mixed performance of the two methods illustrates the importance of users performing a thorough validation in order to determine the influence of simulated precipitation on their chosen impact sector.

1. Introduction

The Intergovernmental Panel on Climate Change (IPCC) has stated in its Fifth Assessment Report that 'it is certain that global mean surface temperature has increased since the late 19th century', with a globally averaged combined ocean and land warming of 0.7-1.1°C from 1880-2012 and 0.5-0.9°C from 1951-2012 (Hartmann et al. 2013). In addition, future temperatures are projected to rise by between 0.3°C and 4.8°C by the end of this century (Collins et al. 2013). Accompanying these rising temperatures is an intensification of the hydrological cycle and the modification of precipitation characteristics, leading to observed and projected increases in the frequency and magnitude of extreme precipitation events such as very intense precipitation and consecutive dry days in many places (Collins et al. 2013;

Hartmann et al. 2013). These changing precipitation characteristics reveal the potential for increasing flooding and drought in the future, bringing about major implications for a wide range of environmental and socio-economic impact sectors including agriculture, landslide risk and soil erosion (Zhang, 2005).

Given these potential implications, assessing the response of a chosen impact sector to changes in future precipitation is an important step in planning future resources and managing hazards. General circulation models (GCMs) are most commonly used to provide the future climate change scenarios necessary for driving impact models. A scale mismatch exists, however, between the spatial resolution at which GCMs provide projections and the much finer resolution at which impact modellers require this information. Downscaling techniques are used to bridge this gap and provide future scenarios at the spatial resolution appropriate for subsequent impact analysis and decision-making. Various downscaling techniques are used for many different impact sectors. Broadly, these approaches can be grouped into either dynamical or statistical downscaling (SD) (Wilby and Dawson 2007).

Dynamical downscaling involves nesting a high-resolution Regional Climate Model (RCM) within a coarser resolution GCM. RCMs provide a spatial resolution of tens of kilometres. Being physically-based, this approach enables small-scale atmospheric features such as low-level jets and orographic precipitation to be better resolved than the host GCM (Wilby and Dawson 2007). The main technical disadvantage is that any biases in the GCM are inherited through the nesting process by which the regional model is developed (Oldfield 2005). For example, gross errors in the precipitation climatology of an RCM may arise if the mid-latitude jet and associated storm tracks are misplaced in the GCM (O'Hare et al. 2005). In addition, although the spatial resolution of RCMs is greatly improved relative to GCMs, direct use of RCM output in impact models is generally discouraged, as suggested by the IPCC guidance for use of RCM output (Mearns et al. 2003). This is firstly because the spatial resolution is still not adequate for various impact sectors relying on site-specific scenarios for point-scale processes, e.g. soil erosion (Mullan et al. 2012a; Mullan 2013). Secondly, RCMs are well known for their systematic errors in predicting daily precipitation, consistently overpredicting the number of wet days and low intensity precipitation yet underestimating intense rainfall (Guo and Senior 2006; Semenov 2007; Maraun et al. 2010; Herrera et al. 2010; Themeßl et al. 2010; Rosenberg et al. 2010; van Roosmalen et al. 2010). One of the key reasons for these shortcomings is the poor representation of convection within parameterisation schemes used in current RCMs (Lenderick et al. 2010). Correction procedures for RCM bias have been widely used to overcome the issues outlined above using model output statistics (MOS) (e.g. Guo and Senior 2006; Schoof et al. 2009; Rosenberg et al. 2010; Themeßl et al. 2010; van Roosmalen et al. 2010). MOS methods can correct RCM precipitation intensity with respect to precipitation amounts and frequency (number of wet days) but cannot modify the temporal sequence of precipitation (Maraun et al. 2010).

SD methods, meanwhile, rely on identifying and developing mathematical transfer functions between observed local climate variables (predictands) and large-scale reanalysis or climate model outputs (predictors) using regression-type methods such as multivariate linear or non-linear regressions (e.g. Corte-Real et al. 1995; Kidson and Thompson 1998; Kilsby et al. 1998; Wilby et al. 1998); principle component analysis (e.g. Karl et al. 1990; Murphy 1999); canonical correlation analysis (e.g. von Storch et al. 1993; Busuioc et al. 1999); principle component analysis (Schubert and Henderson-Sellers, 1997) analogue methods (e.g. Martin et al. 1997; Timbal and McAvaney 2001; Timbal et al. 2003; Zorita and von Storch 1999) kriging; and artificial neural networks (e.g. Trigo and Palutikof 2001; Crane and Hewitson 1998; Wilby et al. 1998). Compared with dynamical downscaling, SD methods are much less computationally demanding and expensive, and can be easily applied to output from many different GCM experiments (Wilby et al. 2004). The major theoretical weakness of SD is that statistical relationships derived for the present day will hold under future climate forcing (Busuioc et al. 1999; Solman and Nuñez 1999; von Storch et al. 2000, Wilby and Wigley 2000, Wilby et al. 2004), i.e. that the climate will remain stationary through time. Predictor estimates and relationships are therefore assumed to be time-invariant, yet it is well recognised that transfer functions may become invalid or weights attached to different predictors could change under future climate forcing (Wilby et al. 2004). Relationships therefore must be critically and carefully assessed as it is not possible to validate future climate conditions with observed records (Arnell et al. 2003).

The above weakness of SD methods is an example of non-stationarity, which describes situations in which the climate system changes through time (Wilby 1998). Non-stationary climates can also represent a problem for SD methods in terms of calibrating models based on time series which change considerably over time. In order to test the robustness of SD methods for simulating non-stationary time series, observed records that exhibit this property can be examined.

2. GPCC vs SDSM

The two contrasting SD techniques used in this paper are both based around transfer function and weather generator approaches. The Generator for Point Climate Change (GPCC) method (Zhang 2005; 2012; Zhang et al. 2012) is a hybrid model combining quantile mapping with a weather generator to develop site-specific climate change scenarios. There are two key downscaling steps in the GPCC process. Firstly, monthly precipitation is spatially downscaled using a quantile mapping method. This involves the development of transfer functions between observed monthly precipitation and reanalysis/model simulated monthly precipitation for a calibration period and a subsequent application of these transfer functions to downscale model simulated monthly precipitation for a future or validation period (Chen et al. 2014a). The second step involves temporally downscaling the spatially downscaled monthly projections to daily data using the weather generator CLIGEN (Nicks and Lane 1989). The key advantage of the GPCC method over many other SD approaches is that it requires monthly rather than daily projections. Monthly projections are generally more accurately simulated than daily projections (Maurer and Hidalgo 2007) and are more readily available from climate models and emissions scenarios (Chen et al. 2014a). In addition, the direct downscaling of precipitation with precipitation as a sole predictor has been found in some cases to capture more explained variance in the predictand than conventional methods that use various other large-scale atmospheric variables (Widmann et al. 2003; Schmidli et al. 2006; Chen et al. 2012a; Chen et al. 2014b). It is also less time consuming than methods that screen and shortlist predictors for model calibration. GPCC has been used and tested extensively for stationary and non-stationary precipitation series across a range of global climatic zones with satisfactory results (Zhang 2005; 2012; Zhang et al. 2012).

The Statistical Downscaling Model (SDSM) (Wilby and Dawson 2007) is frequently described as a hybrid between a regression-based approach and a weather generator, because large-scale daily circulation patterns and atmospheric moisture variables are used to condition local-scale weather generator parameters at individual sites (Wilby and Harris 2006). The underlying philosophy of SDSM relies on the establishment of multiple regressions between station-scale predictands (such as daily rainfall and temperature) and regional-scale predictors (such as mean sea level pressure and near surface vorticity (Wilby and Dawson 2007). The established relationships are then applied to a comparable set of circulation and / or large-scale surface variables simulated by a GCM in order to generate projections of local climate. It is thought that GCMs simulate large-scale atmospheric circulation better than they simulate surface climate variables (Murphy 2000), so in theory the GCM variables applied to SDSM should provide a more realistic basis for downscaling than the sole surface climate variable (precipitation or temperature) applied to GPCC transfer functions. SDSM has been widely used for various impact assessments in 39 countries, yielding over 170 publications (Wilby and Dawson 2013). The model has also been extensively evaluated and performed favourably in model comparison studies for daily precipitation amounts (Khan et al. 2006; Dibike and Coulibaly 2005); precipitation variability (Diaz-Nieto and Wilby 2005); seasonal and annual precipitation totals (Wetterhall et al., 2007a; 2007b); extreme areal average precipitation (Hashmi et al. 2011a); and inter-site correlation of precipitation amounts (Liu et al. 2011) across a range of stationary and non-stationary climates.

Whilst there has been extensive research conducted on comparing dynamical downscaling approaches with statistical downscaling (e.g. Mearns et al. 1999; Murphy 1999; Wilby et al. 2000; Hellstrom et al. 2001; Wood et al. 2004; Haylock et al. 2006; Schmidli et al. 2007), there has been rather less attention afforded to comparing statistical downscaling methods with each other. Wilby et al. (1998) compared a range of weather generator techniques with artificial neural networks (ANNs) for downscaling precipitation across six sites in USA, with the latter performing more poorly owing to failure to adequately simulate wet day occurrence statistics. Zorita and von Storch (1999) compared a simple analogue technique with more complicated SD techniques and found that it simulated winter rainfall for the Iberian Peninsula just as well. Diaz-Nieto and Wilby (2005) compared the change factor (CF) and transfer function-based SD methods for application to low flows in the Thames basin, UK and concluded that transfer function-based SD methods were more appropriate to hydrological impacts modelling since they considered the temporal sequence of precipitation days. These few studies of SD comparisons outlined above generally evaluate simplistic methods against complex techniques, which is probably a consequence of improving techniques with time and the desire for parsimony. In this study, we compare two SD techniques of similar complexities. SDSM has been extensively utilised and evaluated, while GPCC has been less widely utilised but has been established as a competent model across a range of global climatic zones. How the methods compare should therefore be of interest to the SD community. Ultimately both produce site-specific daily series – which is essential for a range of impact sectors including hydrology, soil erosion and crop growth (Zhang 2005). Despite these fundamental similarities, the two techniques differ considerably in terms of data requirements, key model steps, and ultimately yield a different set of advantages and disadvantages for use. These advantages and limitations of GPCC and SDSM are summarised in Table 1. The fact that certain aspects of both models can represent both an advantage and a limitation in certain instances highlights how trade-offs need to made when selecting which SD method to use as no perfect method exists.

This aim of this paper is to compare SDSM and GPCC in terms of their ability to reproduce observed characteristics of non-stationary precipitation series from a range of global climatic zones.

3. Materials & Methods

A general overview of the datasets and methods used for the two models in this study is provided in Table 3.

3.1 Data Sources

3.1.1 Predictands

Observed daily precipitation series were obtained for ten climate stations across the world (Figure 1 and Table 2). Stations were selected on the basis of: 1) completeness of precipitation records to ensure a baseline climatology from 1948 to as close as possible to present (to comply with availability of predictor variables); and 2) a wide geographical spread of stations to capture a diverse range of global climatic zones. The selected stations span four continents and capture precipitation regimes from climatic zones as diverse as the polar arid tundra climate at Resolute Cars, northern Canada, to the humid subtropical climate of Port Macquarie, Australia. Whilst the study would be improved with an examination of further records, the ten stations examined here have been carefully selected to be as representative of the world's precipitation regimes as possible and should therefore facilitate a robust validation of the selected models across a broad range of global climatic zones. The measured daily precipitation series at each station were split into a calibration period and a validation period in a manner that maximised the difference in precipitation between the two periods whilst also ensuring that at least 20 years of the record were retained for the validation period. This ensured the downscaling methods could be tested in non-stationary climates. Relative changes in mean annual precipitation for the validation period relative to the calibration period range from a 21% decrease to a 38% increase.

3.1.2 Predictors

In order to carry out the downscaling analysis using SDSM, daily data were required. A total of 21 large-scale surface and atmospheric predictor variables at a daily temporal resolution were obtained from the National Oceanic and Atmospheric Administration Earth System Research Laboratory Physical Sciences Division. These variables were downloaded for: 1) the grid box directly overlying each of the ten target stations; and 2) an inverse distance weighted (IDW) interpolation of the four adjacent grid boxes positioned closest to the target station. The IDW technique works by predicting new values between the central points of the selected grid squares (in this case four grid squares) within the range of the original values (Burrough and McDonnell, 2004). The advantage of this for climate research is the production of smooth transitions from one grid box to the next rather than abrupt changes which are less realistic in reality. The IDW interpolation technique has been used for smoothing variables between grid boxes on the premise that there is no reduction to the spatial resolution in a range of downscaling studies, e.g. Machguth et al. (2009) and Chen et al. (2014). Use of the inverse distance weighted method allows potential spatial offsets in the predictor-predictand relationship to be examined since neighbouring large scale and surface climate variables from neighbouring grid boxes to the one overlying the target station are considered in the analysis. Reanalysis predictor variables spanning 1948-present with a spatial resolution of 2.5° x 2.5° and representing the 'observed period' were obtained from the National Centre for Environmental Prediction (NCEP). The NCEP Reanalysis project involves the recovery of land surface, ship, radiosonde, aircraft, satellite and other data to assimilate a quality controlled observed record of large-scale circulation variables and surface climate spanning the period from 1948 to present (Kalnay et al. 1996). Extracted predictor variables included geopotential heights, mean air temperature, humidity variables, and a range of secondary airflow variables, all for three atmospheric pressure levels (1000 hPa, 850 hPa and 500 hPa). For the analysis using the GPCC method, monthly precipitation from NCEP representing the 'observed period' was the only data required.

3.2 SDSM Methodology

3.2.1 Predictor screening

All 21 daily predictor variables were examined on a seasonal basis to test their correlation with the full precipitation records at each of the ten stations. The 21 variables were shortlisted to 12 on the basis of those variables exhibiting the strongest correlations with precipitation for each site and season (12 was chosen as this is the maximum number of variables permitted by SDSM for the next step). Subsequently, these 12 variables were further shortlisted to five predictors on the basis of their unique explanatory power, as determined by a partial correlations analysis. The justification for a cut-off at five variables was that the inclusion of additional predictors increases model noise and counters the statistical downscaling ethos of parsimony (e.g. Huth 2005), with five variables evaluated as an appropriate balance between improving model skill and parsimony (Crawford et al. 2007; Mullan et al. 2012b). This generated a statistically "optimum" predictor set for each station and season. This procedure was conducted using predictors from both the overlying grid box and the interpolated grid box, allowing an examination for differences in the optimum predictor sets depending on which grid box was selected. In selecting the grid box to use for downscaling precipitation for each station, the grid box showing higher site-specific values of explained variance relating to the optimum predictor set for that grid box was employed (Table 4).

3.2.2 Model Calibration and Validation

Following selection of the most appropriate grid box, selected predictor variables were then used to calibrate the statistical transfer functions on a monthly basis for each station (Table 5). On the basis of the calibrated monthly models, a weather generator within SDSM was then used to generate precipitation data for the validation period of each station. In the case of wet day occurrence (W_i), there is a direct linear dependency on n predictor variables X_{ij} on day $_i$ (Wilby and Dawson, 2013):

$$W_i = \alpha_0 \sum_{j=1}^{n} \alpha_{jX_{ij}} \qquad (1)$$

under the constraint $0 \leq W_i \leq 1$. Comparison of wet day probability with a random number drawn from a pseudo-random number generator determines whether the day is wet or dry (Wilby et al. 2002). On wet days, precipitation total P_i is calculated using:

$$P_i^k = \beta_0 + \sum_{j=1}^{n} \beta_j X_{ij} + e_i \qquad (2)$$

Where K represents a fourth root transformation designed to make daily wet day amounts match more closely with the normal distribution (Wilby and Dawson 2013). The value of K (0.25) is constrained in such a manner that observed and downscaled precipitation totals are equal for the simulation period (Wilby et al. 1999). Owing to the desire to test the ability of the downscaling techniques in this study under non-stationary conditions. The weather generator produces twenty ensembles of synthetic daily weather series, which helps address uncertainty associated with individual ensemble members (Wilby et al. 2004). All twenty ensembles were stacked together for each station, and the statistics from this compiled record was then compared with the observed precipitation for the same period to enable validation of the model. A similar method for downscaling using SDSM was used in Mullan et al. (2012b).

3.3 GPCC Methodology

3.3.1 Spatial downscaling

Monthly precipitation derived from the NCEP reanalysis was spatially downscaled using a quantile mapping method in two steps. The first step involved establishing the first- and third-order polynomials between observed and NCEP-simulated monthly precipitation quantiles for the calibration period and for all stations. The second step involved using the established polynomials to downscale NCEP-simulated monthly precipitation for the validation period. Since the fitting of the third-order polynomial was consistently better than that of the first-order, the third-order polynomial was used to transform the simulated monthly precipitation values that were within the range in which the third-order polynomial was fitted, while the first-order polynomial was used for the values outside the range (i.e. the linear fit was used for extrapolation). The mean and variance of spatially downscaled monthly precipitation for the validation period were calculated at the target station for further temporal downscaling.

3.3.2 Temporal downscaling

The temporal downscaling involved perturbing CLIGEN parameters based on the spatially downscaled monthly precipitation for the validation period. A first-order, two-state Markov chain is used in CLIGEN to generate precipitation occurrence. The probability of precipitation on a given day is based on the wet or dry status of the previous day, which can be defined in terms of the two conditional transition probabilities: a wet day following a dry day (*P01*) and a wet day following a wet day (*P11*). If a random number drawn from a uniform distribution for each day is less than the precipitation probability for the given previous status, a precipitation event is predicted. For a predicted wet day, a three-parameter skewed normal distribution is used to generate daily precipitation amounts for each month (Nicks and Lane 1989; Nicks et al. 1995). In total, five parameters are needed by CLIGEN to generate daily precipitation series. These include *P11* and *P01* for generating precipitation occurrence, and the mean, standard deviation and skewness coefficient for generating daily precipitation amounts. GPCC only adjusts four parameters and keeps the skewness coefficient unadjusted for the validation period, because there is no easy way to modify the skewness coefficient.

Downscaling of precipitation occurrence involved adjusting three probabilities of precipitation occurrence based on their linear relationships with mean monthly precipitation (R_m). These three probabilities include two conditional transition probabilities (*P11* and *P01*) and one unconditional probability (π). The unconditional probability π can be expressed as:

$$\pi = \frac{P_{01}}{1 + P_{01} - P_{11}} \qquad (3)$$

The adjustment of three probability parameters includes four steps. The first three steps were developed and applied in Zhang (2012) and Zhang et al. (2012), whilst the fourth step was added and applied in Chen et al. (2014). 1) For each month, the observed daily precipitation was divided into two even periods. *P11*, *P01*, π and R_m were respectively calculated for both periods to obtain two data points (one pair for the first period and another for the second period). 2) For each month, the same observed daily precipitation time series was also sorted and divided into wet and dry groups according to the total monthly precipitation. Similarly, *P11*, *P01*, π, and R_m were respectively calculated for both groups to obtain two additional data points (one pair for the wet group and another for the dry group). 3) Linear relationships using linear regression between each of the three probability parameters (dependent) and R_m (predictor) were established using the four data points calculated in step (1) and step (2). The determination coefficient is used as a criterion for selection. 4) For the validation period, the two parameters with the largest coefficient of determination among *P11*, *P01* and π were used for interpolation using the fitted linear equations in step (3) and the spatially downscaled R_m. The remaining parameter was then calculated using equation (3).

The adjusted mean daily precipitation per wet day (u_d) was estimated using equation (4) (Wilks 1992; 1999; Chen et al. 2012b).

$$\mu_d = \frac{\mu_m}{N_d \pi} \qquad (4)$$

where N_d is the number of days in a month and u_m is the mean of spatially downscaled monthly precipitation.

The adjusted daily variance (σ_d^2) was approximated using equation (5), based on the variance of spatially downscaled monthly precipitation (σ_m^2) (Wilks 1992, 1999; Chen et al. 2012b).

$$\sigma_d^2 = \frac{\sigma_m^2}{N_d \pi} - \frac{(1-\pi)(1+r)}{1-r}\mu_d^2 \qquad (5)$$

where r is a dependence parameter defined as:

$$r = P_{11} - P_{01} \qquad (6)$$

All adjusted parameters including *P11*, *P01*, means, and standard deviations of daily precipitation, and the unadjusted skewness of daily precipitation at the calibration period for each month were input to CLIGEN to generate 100 years of daily precipitation for the validation period. CLIGEN-generated time series for the validation period were then compared with SDSM-generated and observed data for the same period.

3.4 Statistical Analysis

An overview of the statistical approach to validating GPCC and SDSM against observed precipitation for the validation period is given in Table 6. These statistics were calculated for four temporal resolutions: mean daily precipitation (i.e. mean of all summed days in the record), mean monthly precipitation (i.e. mean of all summed months in the record), mean annual precipitation (i.e. mean of all summed years in the record), and annual maximum daily precipitation (i.e. mean of maximum daily precipitation value for each year). In addition, the temporal structure of the two downscaling methods was evaluated with respect to its ability to reproduce dry and wet spells by plotting the cumulative frequencies of observed and downscaled dry and wet spell lengths.

4. Results

Results showing the ability of the two downscaling techniques to replicate various characteristics of precipitation for the ten climate stations analysed in this study are presented and discussed in this section. Tables 7-10 display observed precipitation amounts and RE of both downscaling methods for each station and statistic at each of the four temporal resolutions respectively as outlined in the Methods section and shown in Table 6. Also shown in these tables is the mean RE and mean ARE of each downscaling method across all ten stations for all statistics. It should be pointed out that the observed validation periods are 20 years for most stations while the simulated data durations are 100 years for GPCC and 20 years for SDSM. Their direct comparisons for the extreme events such as the 'all time' maximum are crude and only have limited values in some cases.

4.1 Mean Daily Precipitation (MDP)

For most of the statistics, there is close agreement between observed precipitation and precipitation simulated by the two downscaling techniques. In particular, the mean, standard deviation and percentiles are generally well simulated. As shown in Table 7, the mean ARE for the mean of MDP across all stations is 10.7% and 8.4% respectively for GPCC and SDSM, which is reasonably close to the observed mean. Despite the relatively low mean ARE, GPCC underestimates the mean by as much as 26% at the low precipitation station of Resolute Cars and by 21% at the very wet station of Cataract Dam, whilst SDSM overestimates by as much as 16% at the very wet station of Fort Pierce. This indicates that while both techniques simulate the mean reasonably well, in many instances they do not perform as well for those stations with a more extreme mean daily precipitation. The mean RE of -8.5% for GPCC and 0.1% for SDSM reveals the underestimating bias of GPCC and the mixed bias of SDSM.

The mean ARE for the standard deviation is 15% and 21% for GPCC and SDSM respectively. Generally, GPCC overestimates the standard deviation of daily precipitation (at seven stations – mean RE of 4.6%), while SDSM underestimates at nine stations with a mean RE of -13.3%. This indicates that the spread of values across the extremes should be lower for SDSM than GPCC, meaning the former is likely to overestimate lower precipitation amounts and underestimate higher precipitation amounts, with the reverse likely true of the latter.

This trend can be picked up when examining the percentiles. For lower precipitation amounts (Q25), GPCC underestimates at nine stations (mean RE of -32%) whilst SDSM overestimates at eight stations (mean RE of 44.1%), with GPCC overestimating at five stations for Q99 (mean RE of 5.1%) and SDSM underestimating at eight of them (mean RE of -12.4%). In keeping with overestimating the upper extremes, GPCC overestimates the maximum of MDP at nine stations, with a mean RE of 56%. Yet, despite largely underestimating Q99, SDSM overestimates the maximum at six stations, with a mean RE of 27%.

Neither model simulates skewness well. GPCC largely overestimates (at eight stations with a mean RE of 24.1%) whilst SDSM largely underestimates (at eight stations with a mean RE of -12.9%), which is in keeping with their treatment of Q99.

The treatment of the mean number of wet days is generally better for SDSM than GPCC, reflected by the lower mean ARE in the former (7.1% as opposed to 11.9% respectively). GPCC overestimates this statistic at nine stations with a mean RE of 9.6%, whilst SDSM underestimates at seven with a mean RE of -2.8%.

4.2 Mean Monthly Precipitation (MMP)

The agreement between observed and simulated precipitation is very similar to that of MDP for most statistics, but the sign of the error is somewhat different, as is the greatly reduced number of stations where certain percentiles are seriously under or overestimated. As shown in Table 8, the mean ARE across all stations is 10.2% and 8.4% for GPCC and SDSM respectively, with REs for individual stations generally reduced compared with MDP. Despite this improvement in REs over MDP, there is one large exception for both models, as GPCC overestimates the mean by up to 35.2% for the very wet station of Port Macquarie and SDSM underestimates the mean by up to 25.2% for the very dry station of Resolute Cars. Again, this reflects the difficulty of simulation for extreme stations. Nonetheless, other extreme stations are well simulated by both models for the mean.

Standard deviation is better simulated by GPCC than SDSM (mean ARE of 14.4% for GPCC as opposed to 32% by SDSM). This time, both models underestimate standard deviation at more stations (seven for GPCC with a mean RE of -4.3% and nine for SDSM with a mean RE of -4.2%), yet there is one massive overestimation of 139% by SDSM at the wet station of Campinas. In theory, therefore, both models should overestimate lower extremes and underestimate the upper extremes (notwithstanding stations that overestimate the standard deviation).

This trend is visible when examining the percentiles. Low precipitation amounts (Q25) are overestimated by both models at seven out of the ten stations, with a mean RE of 14.6% and 2.1% for GPCC and SDSM respectively. High precipitation amounts (Q99) are underestimated by both models at eight out of the ten stations (mean RE of -2.1% for GPCC and -3.1% for SDSM), yet both models overestimate at Ottawa and one more of the wettest stations (Port Macquarie and Campinas respectively) – mostly stations that overestimated the standard deviation. This again reflects how the simulation of standard deviation is a good indicator of how the extremes will be simulated. Despite this relationship, the maximum for MMP is overestimated by both models, at ten stations with a mean RE of 31.4% for GPCC and at nine stations with a mean RE of 39.8% for SDSM.

The skewness coefficient may be responsible for this, as it is overestimated by GPCC at eight stations (mean RE = 38.9%) and overestimated by SDSM at five stations (mean RE = 17.1%).

Zhang et al. (2012) evaluated the ability of GPCC in downscaling monthly precipitation to daily series at the same ten stations in this study without the spatial downscaling step. Monthly precipitation at these stations was directly used in GPCC for the temporal disaggregation. Their results showed that GPCC preserved and reproduced monthly statistics including mean, standard deviation, skewness, and percentiles very well. The less satisfactory performance found in this work indicates that errors in fitting the transfer functions for spatial downscaling as well as in NCEP-simulated monthly precipitation for the validation period might have affected the downscaling results.

4.3 Mean Annual Precipitation (MAP)

The mean ARE is identical to that of MMP for the mean at 10.2% and 8.4% respectively for GPCC and SDSM, as is the RE for individual stations, all of which indicates that the mean for MAP is simulated reasonably well by both models (with the same exceptions as for MMP).

As was the case with MMP, the standard deviation is underestimated at most stations by both models (eight stations in the case of GPCC with a mean RE of -10%), and nine in the case of SDSM with a mean RE of -15.9%.

This time, however, the expected response in extremes does not quite hold true. Both models overestimate Q25 at only half the stations (mean RE of 3.7% for GPCC and 0.7% for SDSM), though the overestimations are much higher than the underestimations at the other half (e.g. overestimations up to 38.2% at the wet station of Port Macquarie for GPCC). Underestimations of the upper percentile (Q99) and maximum, as might be expected with a low standard deviation, occurs at just four stations For GPCC and just three for SDSM, with large overestimations of up to 37.9% by GPCC for Brenham.

Again, the skewness coefficient can help explain why these higher precipitation amounts are projected despite a lower standard deviation. The skewness coefficient is overestimated at many of the same stations that Q99 and the maximum are overestimated for, which again demonstrates the role skewness plays in generating extreme precipitation amounts.

4.4 Annual Maximum Daily Precipitation (AMDP)

Table 10 shows the mean ARE for GPCC and SDSM is 18.4% and 23.4% respectively for the mean, which is approximately double the mean ARE than any of the other temporal resolutions.

The mean is overestimated at six stations by GPCC (mean RE of 12.5%) and underestimated at eight stations by SDSM (mean RE of -15.2%). Since we are dealing with extremes, this is to be expected.

The standard deviation is overestimated at seven stations by GPCC (mean RE of 24.9%) and underestimated for eight stations by SDSM (mean RE of -20.8%). Once again, this influence comes through in the percentiles, with Q99 overestimated at eight stations by GPCC and underestimated at six stations by SDSM, with a mean RE of 44.7% and -5.7% respectively. There is less evidence of the link between standard deviation and precipitation extremes from the lower percentiles (Q25) as GPCC underestimates at only half the stations (mean RE of 10%) and SDSM overestimates for only two (mean RE of -13%). This illustrates that GPCC provides a wider spread of values across the extremes, which is reflected by the generally higher standard deviation for GPCC.

Skewness is overestimated at six stations by GPCC (mean RE of 404.3%) and SDSM (mean RE of 499.4% and an exceptionally high RE of 4419.4% at Barkerville) which helps explain the overestimation of the maximum by both models (mean RE of 56% for GPCC and 27% by SDSM).

4.5 Dry and Wet Spell Lengths

The temporal structure of GPCC- and SDSM-generated daily precipitation is evaluated with respect to reproducing the dry and wet spells. The cumulative frequencies of dry and wet spells generated by GPCC and SDSM for the validation period are compared with those directly calculated from the observed precipitation of the same period for all 10 stations (Figures 2 and 3).

Overall, SDSM overestimates the frequencies of both dry and wet periods, especially for short dry and wet spells, indicating that SDSM generates too many continuously short dry and wet events. Similar results were also found by Chen et al. (2012a) in their study. GPCC performs much better than SDSM for downscaling distributions of both wet and dry spells, even though the dry and wet spells can be slightly overestimated or underestimated for some stations. However, GPCC overestimates the longest dry and wet spells for eight stations respectively (Table 11). In contrast, SDSM underestimates the longest dry and wet spells for four and eight stations respectively, as also shown in Table 11. Both models show a better performance for downscaling wet spells than dry spells, especially for SDSM.

5. Discussion

Both the GPCC and SDSM models can in many instances closely reproduce a range of observed characteristics of precipitation for non-stationary global climates, but there are also considerable deviations for certain statistics at certain temporal resolutions. Some potential explanations for these factors, based on the workings of the two models and the input data used to drive them, are considered in this section.

5.1 Non-stationarity

A key factor responsible for differences between observed and simulated precipitation characteristics (for all statistics and temporal resolutions) is the issue of non-stationarity. Although this study aims to test if two downscaling methods can reproduce closely characteristics of observed precipitation under non-stationary climates, it is to be expected that regression weights will change through time and result in underestimations and overestimations during the validation period (Wilby et al. 2004). This major theoretical weakness of SD is well known, and requires careful screening of appropriate predictor variables to guard against the ‘time invariance’ assumption (Arnell et al. 2003). Precipitation amounts are prescribed during the calibration procedure, but since the calibration and validation periods were selected to maximise the difference in mean annual precipitation between them, it is to be expected that the application of transfer functions developed for the calibration period to the validation period will result in small differences between observed and simulated means and distribution statistics. It is difficult to attribute this cause of error to specific distribution statistics, but there is little doubt that this is a factor causing some of the simulation error. These deviations are also simulated in Zhang (2012), Zhang et al. (2012) and Chen et al. (2014a).

5.2 NCEP biases

In validation studies of NCEP, significant regional biases have been found between both reanalyses and observations (e.g. Higgins et al. 1996; Mo and Higgins, 1996; Widmann and Bretherton, 1999). In this respect, any under or overestimation in NCEP precipitation for the validation period compared with the calibration period will lead to a similar prediction in the downscaling models. This is likely to be one of the reasons for the differences between observed and simulated precipitation for both methods. Zhang et al. (2012) concluded this was likely one of the causes of simulation error based on their study of the same ten stations used here.

As both methods rely on NCEP data in model calibration, Both methods are subject to biases from NCEP. The direction and magnitude of those biases, however, will be inherently different owing to the fact that GPCC downscales from NCEP simulated surface precipitation at a monthly temporal resolution, as opposed to the use of NCEP simulated large-scale predictor variables at a daily temporal resolution in SDSM. Differences in the temporal resolution and skill in simulating the different NCEP variables will undoubtedly be one of the factors causing the differences in the direction and magnitude of simulated biases. Generally, monthly simulations are thought to be more skilfully simulated than daily variables (Maurer and Hidalgo, 2007), but surface variables are less well simulated than the large-scale variables (Murphy, 2000). Once again, however, it is difficult to pinpoint what specific distribution statistics these differences impact most. This highlights that GPCC and SDSM both have advantages and disadvantages based on the nature of the input data alone.

5.3 Model Differences

In addition to the non-model based factors outlined above, the different downscaling steps in each of the methods may be a key factor impacting the results. The weather generator in SDSM produces daily series based on regression models developed at a monthly temporal resolution. Precipitation amounts and the temporal sequence of precipitation are both derived from the same monthly regression models. This approach does not facilitate the explicit downscaling of these transition probabilities in the same manner as for GPCC, as the transition probabilities are downscaled implicitly in the same step as precipitation amounts during calibration. The use of the unconditional precipitation occurrence probability of Equation 1 without explicitly simulating wet-following-wet and wet-following-dry day probability as in GPCC limits the ability of SDSM to accurately simulate the distributions of wet and dry spells. The use of the linear regression of equation 2 to simulate daily precipitation amounts has an inherent tendency to overestimate small amounts (events) and underestimate large amounts (events). Nearing (1998) has reported that all simulation models including regression models are intended to predict mean values, which would overestimate lower values and underestimate large values. This may be one of the reasons why SDSM overestimates the low precipitation amounts and underestimates the large events, and it may indicate that bias correction is more necessary for SDSM. It is postulated that the use of the bias correction setting within SDSM may not be well placed to address this issue in any case because one correction factor cannot correct both overestimation for small storms and underestimation for large storms. Since SDSM is calibrated on a monthly basis, one single empirically derived bias correction ratio is applied to each monthly model, and this correction ratio is constrained to equalise observed and simulated precipitation totals for the calibration period (Wilby et al. 1999). Under non-stationary conditions, which the stations in this study are all subject to, the constraint applied to the correction factor when developing the transfer functions for the calibration period is likely to underestimate those larger events that may occur outside the range of observations during the validation period. In this respect, the SDSM bias correction ratio may be inadequate to correct precipitation amounts of the largest events, and may be too large to correct the smaller events. The lack of spread in generated daily precipitation amounts with SDSM may be because the probability distribution function is not used in daily precipitation generation. That is, the distribution parameters such as standard deviation are not explicitly used in the generation process.

In the case of GPCC, bias correction is inherent in the spatial dowscaling steps where quantile mapping is used to adjust the distribution of NCEP simulated precipitation. GPCC may be better placed to simulate the two stage conditional processes of precipitation (occurrence and amount) due to the explicit spatiotemporal downscaling approaches used. The explicit treatment of spatiotemporal variability by GPCC mentioned above is likely to be the reason why it better simulates wet and dry spell lengths. As transition probabilities are downscaled to daily series from mean monthly precipitation in a series of explicit steps, the wet-following-wet day probability, wet-following-dry day probability, means and variances are explicitly treated to fully represent the temporal structure of precipitation and precipitation distribution of daily amounts. Zhang (2007) highlights the more appropriate role of this explicit approach compared with an implicit approach without separate spatial and temporal downscaling steps for downscaling the temporal sequence of precipitation and their extremes. In GPCC, probability distribution fitting from a skewed normal distribution is used to generate precipitation amounts, in which daily precipitation variance is downscaled and directly used in the generation. Unlike SDSM, use of these probability distributions allows the generation of new extreme values outside the range of observations and this may be why large events are overestimated. Also, because GPCC generated 100 years of data compared to the 20 year observed record, this time mismatch is expected to provide greater extremes in GPCC – thus comparisons of extremes for GPCC should be seen as crude and preliminary.

6. Conclusions and Implications

The generation of realistic future precipitation scenarios is crucial to impact modelling and subsequent resource and hazard planning for a wide variety of environmental and socio-economic impact sectors. This study sought to test two different statistical downscaling methods in terms of their ability to reproduce observed characteristics of precipitation at a range of temporal resolutions for ten non-stationary climates across the world. The following key conclusions can be drawn from this study:

- Both the GPCC and SDSM models can reproduce mean precipitation amounts with a reasonable degree of similarity to the observed mean for MDP, MMP and MAP, with a mean ARE across all stations of close to 10% in all cases. Non-stationarities between the calibration and validation period and/or biases in NCEP simulation are likely responsible for the differences in many cases.
- Relative Errors are much larger for AMDP. GPCC overestimates at most stations (up to 60%), whilst SDSM underestimates at most stations (by up to 59%). This indicates that GPCC may overestimate extreme values of precipitation, whilst SDSM is more

likely to underestimate these. This is likely to be related to the fitting of probability distributions of daily precipitation in GPCC in overestimating extremes, and possibly the fact that SDSM does not downscale based on probability distributions of precipitation.

- Simulation of standard deviation is closely tied up with the simulation of both low and high extremes. Standard deviation tends to be overestimated by GPCC in many cases, which stretches the precipitation values across the percentiles and results in underestimation of low precipitation amounts (Q25) and overestimation of high precipitation amounts (Q99 and maximum). The reverse is true for SDSM with an underestimated standard deviation resulting in overestimated lower precipitation extremes and underestimated upper extremes.
- In cases where standard deviation cannot explain the RE in the extremes, the skewness coefficient may play a key role. Skewness is generally underestimated by SDSM, which results in underestimated upper extremes, whilst GPCC tends to overestimate skewness and thus also overestimate maximum precipitation amounts.
- SDSM tends to overestimate wet and dry spell frequency, whilst GPCC generally simulated these more closely to the observed temporal structure. This is likely to be related to the explicit spatiotemporal downscaling of transition probabilities in GPCC. This may make GPCC more appropriate to those impact sectors where the temporal sequence of precipitation events is critical, e.g. hydrology.
- Most of this evidence points towards the likelihood that GPCC is more likely to overestimate precipitation extremes and thereby overestimate the effects on whatever impact sector is being simulated, whilst SDSM is likely to do the opposite and underestimate the impacts.
- The study reveals the importance of performing a thorough validation of downscaled precipitation scenarios in order to consider the reliability of modelled scenarios of a particular impact sector in response to climate change.

Acknowlegments

The authors wish to thank Dr Bofu Yu from Griffith University, Australia, for providing daily precipitation data for Cataract Dam and Port Macquarie stations and to Dr Alfredo Borges de Campos from Universidade Federal de Goias, Brazil, for precipitation data for Campinas station. Thanks must also go to Prof. Rob Wilby and Dr Christian Dawson for use of their SDSM model and to NOAA for the availability of predictor variables.

References

Arnell NW, Hudson DA, Jones RG. 2003. Climate change scenarios from a regional climate model: Estimating change in runoff in southern Africa. Journal of Geophysical Research 108: 1 – 17.

Burrough PA, McDonnell RA. 1998. Principles of Geographical Information Systems. Oxford University Press.

Busuioc A, von Storch H, Schnur R. 1999. Verification of GCM-generated regional seasonal precipitation for current climate and of statistical downscaling estimates under changing climate conditions. Journal of Climate 12: 258 – 272.

Chen J, Brissette FP, Leconte R. 2012a. Coupling statistical and dynamical methods for spatial downscaling of precipitation. Climatic Change 114: 509 – 526.

Chen J, Brissette FP, Leconte R. 2012b. Downscaling of weather generator parameters to quantify the hydrological impacts of climate change. Climate Research 51(3): 185 – 200.

Chen J, Brissette FP, Leconte R. 2014b. Assessing regression-based statistical approaches for downscaling precipitation over North America. Hydrological Processes 28: 3482–3504.

Chen J, Zhang XC, Brissette FP. 2014a. Assessing scale effects for statistically downscaling precipitation with GPCC model. International Journal of Climatology 34: 708 – 727.

Collins M, Knutti R, Arblaster J, Dufresne J-L, Fichefet T, Friedlingstein P, Gao X, Gutowski WJ, Johns T, Krinner G, Shongwe M, Tebaldi C, Weaver AJ, Wehner M. 2013. Long-term Climate Change: Projections, Commitments and Irreversibility. In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker TF, Qin D, Plattner G-K, Tignor M, Allen SK, Boschung J, Nauels A, Xia Y, Bex V, Midgley PM. (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Corte-Real J, Zhang X, Wang X. 1995. Downscaling GCM information to regional scales: a non-parametric multivariate regression approach. Climate Dynamics 11: 413—424

Crane RG, Hewitson BC. 1998. Doubled CO2 precipitation changes for the Susquehanna Basin: downscaling from the GENESIS General Circulation Model. International Journal of Climatology 18: 65–76.

Crawford T, Betts NL, Favis-Mortlock DT. 2007. GCM grid-box choice and predictor selection associated with statistical downscaling of daily precipitation over Northern Ireland. Climate Research 34: 145 – 160.

Diaz-Nieto J, Wilby RL. 2005. A comparison of statistical downscaling and climate change factor methods: impacts of low flows in the river Thames, United Kingdom. Climatic Change 69: 245-268.

Dibike YB, Coulibaly P. 2005. Hydrological impact of climate change in the Saguenay watershed: comparison of downscaling methods and hydrologic models. Journal of Hydrology 307, 145-163.

Flanagan DC, Nearing MA. 1995. USDA - Water Erosion Prediction Project (WEPP) Hillslope Profile and Watershed Model Documentation. West Lafayette, IN., USA. National Soil Erosion Research Laboratory, USDA - Agricultural Research Service.

Guo Y, Senior MJ. 2006. Climate model simulation of point rainfall frequency characteristics. Journal of Hydrologic Engineering. DOI: 10.1061/(ASCE)1084-0699.

Hartmann DL, Klein Tank AMG, Rusticucci M, Alexander LV, Brönnimann S, Charabi Y, Dentener FJ, Dlugokencky EJ, Easterling DR, Kaplan A, Soden BJ, Thorne PW, Wild M, Zhai PM. 2013. Observations: Atmosphere and Surface. In: Climate Change 2013: The Physical Science Basis. Contribution of Working Group I to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change [Stocker TF, Qin D, Plattner G-K, Tignor M, Allen SK, Boschung J, Nauels A, Xia Y, Bex V, Midgley PM. (eds.)]. Cambridge University Press, Cambridge, United Kingdom and New York, NY, USA.

Haylock, MR, Cawley GC, Harpham C, Wilby RL, Goodess CM. 2006. Downscaling heavy precipitation over the United Kingdom: a comparison of dynamical and statistical methods and their future scenarios. International Journal of Climatology 26: 1397-1415.

Hellström C, Chen D, Achberger C, Räisänen J. 2001. Comparison of climate change scenarios for Sweden based on statistical and dynamical downscaling of monthly precipitation. Climate Research 19: 45-55.

Herrera S, Fita L, Fernandez J, Gutierrez JM. 2010. Evaluation of the mean and extreme precipitation regimes from the ENSEMBLES regional climate multimodel simulations over Spain. Journal of Geophysics Research. DOI: 10.1029/2010JD013936.

Higgins RW, Mo KC, Schubert SC. 1996. The moisture budget of the central United States as evaluated in the NCEP/NCAR and the NASA/DAO reanalyses. Mon. Wea. Rev. 124: 939–963.

Huth R. 2005. Downscaling of humidity variables: a search for suitable predictors and predictands. International Journal of Climatology 25: 243 – 250.

Kalnay E, Kanamitsu M, Kistler R, Collins W, Deaven D, Gandin L, Iredell M, Saha S, White G, Woollen J, Zhu Y, Leetmaa A, Reynolds B, Chelliah M, Ebisuzaki W, Higgins W, Janowiak J, Mo KC, Ropelewski C, Wang J, Jenne, R, Joseph D. 1996. The NCEP/NCAR 40-year reanalysis project. Bul. Am. Meteor. Soc. 77: pp. 437 – 471.

Karl TR, Wang WC, Schlesinger ME, Knight RW, Portman D. 1990. A method of relating General Circulation Model simulated climate to the observed local climate. Part I: seasonal statistics. Journal of Climate 3: 1053–1079.

Khan MS, Coulibaly P, Dibike Y. 2006. Uncertainty analysis of statistical downscaling methods. Journal of Hydrology 319, 357-382.

Kidson JW, Thompson CS. 1998. A comparison of statistical and model-based downscaling techniques for estimating local climate variations. Journal of Climate 11: 735–753.

Kilsby CG, Cowpertwait PSP, O'Connell PE, Jones PD. 1998. Predicting rainfall statistics in England and Wales using atmospheric circulation variables. International Journal of Climatology 18: 523–539.

Lenderink G. 2010. Exploring metrics of extreme daily precipitation in a large ensemble of regional climate model simulations. Climate Research 44: 151–166.

Machguth H, Paul F, Kotlarski S, Hoelze M. 2009. Calculating distributed glacier mass balance for the Swiss Alps from regional climate model output: a methodical description and interpretation of the results. Journal of Geophysical Research Atmospheres 114, DOI: 10.1029/2009JD011775.

Maraun D, Wetterhall F, Ireson AM, Chandler RE, Kendon EJ, Widmann M, Brienen S, Rust HW, Sauter T, Themebl M, Venema VKC, Chun KP, Goodess CM, Jones RG, Onof C, Vrac M, Thiele-Eich L. 2010. Precipitation downscaling under climate change. Recent developments to bridge the gap between dynamical models and the end user. Reviews of Geophysics. DOI: 10.1029/2009RG000314.

Martin E, Timbal B, Brun E 1997. Downscaling of general circulation model outputs: simulation of the snow climatology of the French Alps and sensitivity to climate change. Climate Dynamics 13: 45–56.

Maurer EP, Hidalgo HG. 2007. Utility of daily vs. monthly large-scale climate data: an intercomparison of two statistical downscaling methods. Hydrology and Earth System Sciences Discussions 4: 3413 – 3440.

Mearns LO, Giorgi F, Whetton P, Pabon D, Hulme M, Lal M. 2003. Guidlines for use of climate scenarios developed from Regional Climate Model experiments. Technical Report. The IPCC Data Distribution Centre, Norwich, UK, 38.

Mo KC, Higgins RW. 1996. Large-scale atmospheric moisture transport as evaluated in the NCEP/NCAR and the NASA/ DAO reanalyses. Journal of Climate 9: 1531–1545.

Mullan DJ. 2013. Soil erosion under the impacts of future climate change: assessing the statistical significance of future changes and the potential on-site and off-site problems. Catena, vol. 109, 234-246.

Mullan DJ, Favis-Mortlock DT, Fealy R. 2012a. Addressing key limitations associated with modelling soil erosion under the impacts of future climate change. Agricultural and Forest Meteorology 156: 18 – 30.

Mullan DJ, Fealy R, Favis-Mortlock DT. 2012b. Developing site-specific future temperature scenarios for Northern Ireland: addressing key issues employing a statistical downscaling approach. International Journal of Climatology 32(13): 2007-2019.

Murphy J. 1999. An evaluation of statistical and dynamical techniques for downscaling local climate. Journal of Climate 12: 2256–2284.

Murphy J. 2000. Predictions of Climate Change over Europe using Statistical and Dynamical Downscaling techniques. International Journal of Climatology 20: 489 – 501.

Nearing MA. 1998. Why soil erosion models over-predict small soil losses and under-predict large soil losses. Catena 32: 15-22.

Nicks AD, Lane LJ, Gander GA. 1995. Chapter 2: Weather generator. In USDA−Water Erosion Prediction Project: Hillslope Profile and Watershed Model Documentation, Flanagan DC, Nearing MA (eds) NSERL Report No. 10. USDA-ARS National Soil Erosion Research Laboratory: West Lafayette, IN.

Nicks AD, Lane LJ. 1989. Chapter 2: Weather generator. In USDAWater Erosion Prediction Project: Hillslope Profile Version, Lane LJ, Nearing MA (eds) NSERL Report No. 2. USDA-ARS National Soil Erosion Research Laboratory: West Lafayette, IN.

O'Hare G, Sweeney J, Wilby RL. 2005. Weather, climate, and climate change: human perspectives. New York: Prentice Hall.

Oldfield F. 2005. Environmental Change: Key Issues and Alternative Approaches. Cambridge: Cambridge University Press.

Rosenberg EA, Keys PW, Booth DB, Harley D, Burkey J, Steinemann AC, Lettenmaier DP. 2010. Precipitation extremes and the impacts of climate change on stormwater infrastructure in Washington State. Climatic Change 102: 319–349.

Schmidli J, Frei C, Vidale PL. 2006. Downscaling from GCM precipitation: a benchmark for dynamical and statistical downscaling methods. International Journal of Climatology 26: 679 – 689.

Schmidli J, Goodess CM, Frei C, Haylock MR, Hundecha Y, Ribalaygua J, Schmith T. 2007. Statistical and dynamical downscaling of precipitation: an evaluation and comparison of scenarios for the European Alps. Journal of Geophysical Research Atmospheres 112: DOI: 10.1029/2005JDO07026

Schoof JT, Shin DW, Cocke S, LaRow TE, Lim Y-K, O'Brien JJ. 2009. Dynamically and statistically downscaled seasonal temperature and precipitation hindcast ensembles for the southeastern USA. International Journal of Climatology 29: 243–257.

Schubert S, Henderson-Sellers A. 1997. A statistical model to downscale local daily temperature extremes from synoptic-scale atmospheric circulation patterns in the Australian region. Climate Dynamics 13: 223-234.

Semenov MA. 2007. Development of high-resolution UKCIP02-based climate change scenarios in the UK. Agricultural and Forest Meteorology 144: 127–138.

Solman S and Nuñez M. 1999. Local estimates of global climate change: a statistical downscaling approach. International Journal of Climatology 19: 835 – 861.

Themeßl MJ, Gobiet A, Leuprecht A. 2010. Empirical statistical downscaling and error correction of daily precipitation from regional climate models. International Journal of Climatology. DOI: 10.1002.joc.2168.

Timbal B, Dufour A, McAvaney B. 2003. An estimate of future climate change for western France using a statistical downscaling technique.

Timbal B, McAvaney BJ. 2001. An analogue-based method to downscale surface air temperature: application for Australia. Climate Dynamics 17: 947–963

Trigo RM, Palutikof JP. 2001. Precipitation scenario over Iberia: a comparison between direct GCM output and different downscaling techniques. Journal of Climate 14: 4422–4446.

van Roosmalen L, Christensen JH, Butts MB, Jensen KH, Refsgaard JC. 2010. An intercomparison of regional climate model data for hydrological impact studies in Denmark. Journal of Hydrology 380: 406–419.

von Storch H, Hewitson B, Mearns L. 2000. Review of empirical downscaling techniques. Regional climate development under global warming. Iversen T, Hoiskar BAK (eds). General Technical Report No. 4. Conf. Proceedings RegClim Spring Meeting Jevnaker, Torbjornrud, Norway, 29–46.

Von Storch H, Zorita E, Cubasch U. 1993. Downscaling of global climate change estimates to regional scales: An application to Iberian rainfall in wintertime. Journal of Climate 6: 1161 – 1171.

Wetterhall F, Bárdossy A, Chen D, Halldin S, Xu C. 2007a. Daily precipitation-downscaling techniques in three Chinese regions. Water Resources Research 42, W11423, DOI: 10.1029/2005WR004573.

Wetterhall, F, Halldin S, Xu CY. 2007b. Seasonality properties of four statistical-downscaling methods in central Sweden. Theoretical and Applied Climatology 87, 123-137.

Widmann M, Bretherton CS. 2000. Validation of mesoscale precipitation in the NCEP Reanalysis using a new gridcell dataset for the Northwestern United States.

Widmann M, Bretherton CS, Salathé Jr EP. 2003. Statistical precipitation downscaling over the Northwestern United States using numerically simulated precipitation as a predictor. International Journal of Climatology 16(5): 799 – 816. Journal of Climate 13: 1936-1950.

Wilby, RL. 1998. Statistical downscaling of daily precipitation using daily airflow and teleconnection indices. Climate Research 10: 163 – 178.

Wilby RL, Charles SP, Zorita E, Timbal B, Whetton P, Mearns OL. 2004. Guidelines for the use of Climate scenarios developed from Statistical downscaling methods. Available at http://ipcc-ddc.cru.uea.ac.uk/guidelines/dgm_no2_v1_09_ 2004. pdf.

Wilby RL, Dawson CW. 2007. SDSM 4.2- A decision support tool for the assessment of regional climate change impacts, Version 4.2 User Manual. Lancaster University: Lancaster / Environment Agency of England and Wales.

Wilby RL, Dawson CW. 2013. The Statistical DownScaling Model: insights from one decade of application. International Journal of Climatology 33(7): 1707 – 1719.

Wilby RL, Dawson CW, Barrow EM. 2002. SDSM-A decision support tool for the assessment of regional climate change impacts. Environmental Modelling & Software 17: 145 – 157.

Wilby RL, Harris I. 2006. A framework for assessing uncertainties in climate change impacts: low flow scenarios for the River Thames, UK. Water Resources Research 42 (2): W02419.1-W02419.10.

Wilby RL, Hay LE, Leavesley GH. 1999. A comparison of downscaled and raw GCM output: implications for climate change scenarios in the San Juan river basin, Colorado. Journal of Hydrology 225: 67-91.

Wilby RL, Wigley TML. 2000. Precipitation predictors for downscaling: Observed and general circulation model relationships. International Journal of Climatology 20: 641 – 661.

Wilby RL, Wigley TML, Conway D, Jones PD, Hewitson BC, Main J, Wilks DS. 1998b. Statistical downscaling of general circulation model output: a comparison of methods. Water Resources Research 34: 2995–3008.

Wilks DS. 1992. Adapting stochastic weather generation algorithms for climate change studies. Climatic Change 22: 67 – 84.

Wilks DS. 1999. Multisite downscaling of daily precipitation with a stochastic weather generator. Climate Research 11: 125 – 136.

Wood AW, Leung LR, Sridhar V, Lettenmaier DP. 2004. Hydrologic implications of dynamical and statistical approaches to downscaling climate model outputs. Climatic Change 62: 189-216.

Zhang X-C. 2007. A comparison of explicit and implicit spatial downscaling of GCM output for soil erosion and crop production assessments. Climatic Change 84: 337-363.

Zhang X-C. 2005. Spatial downscaling of global climate model output for site-specific assessment of crop production and soil erosion. Agricultural and Forest Meteorology 135: 215 – 229.

Zhang X-C. 2012. Verifying a temporal disaggregation method for generating daily precipitation of potentially non-stationary climate change for site-specific impact assessment. International Journal of Climatology 33: 326 – 342.

Zhang X-C, Chen J, Garbrecht JD, Brissette FP. 2012. Evaluation of a weather generator-based method for statistically downscaling non-stationary climate scenarios for impact assessment at a point scale. Transactions of the ASABE 55(5): 1 – 12.

Zorita E, von Storch H. 1999. The analog method as a simple statistical downscaling technique: comparison with more complicated methods. Journal of Climate 12: 2474-2489.

Figure 1

Click here to download Figure: Figure1.tif

Figure 1 caption

Figure 1. Location of the ten climate stations used in this study. Details for the stations are provided in Table 2.

Figure 2
Click here to download Figure: Figure2.tif

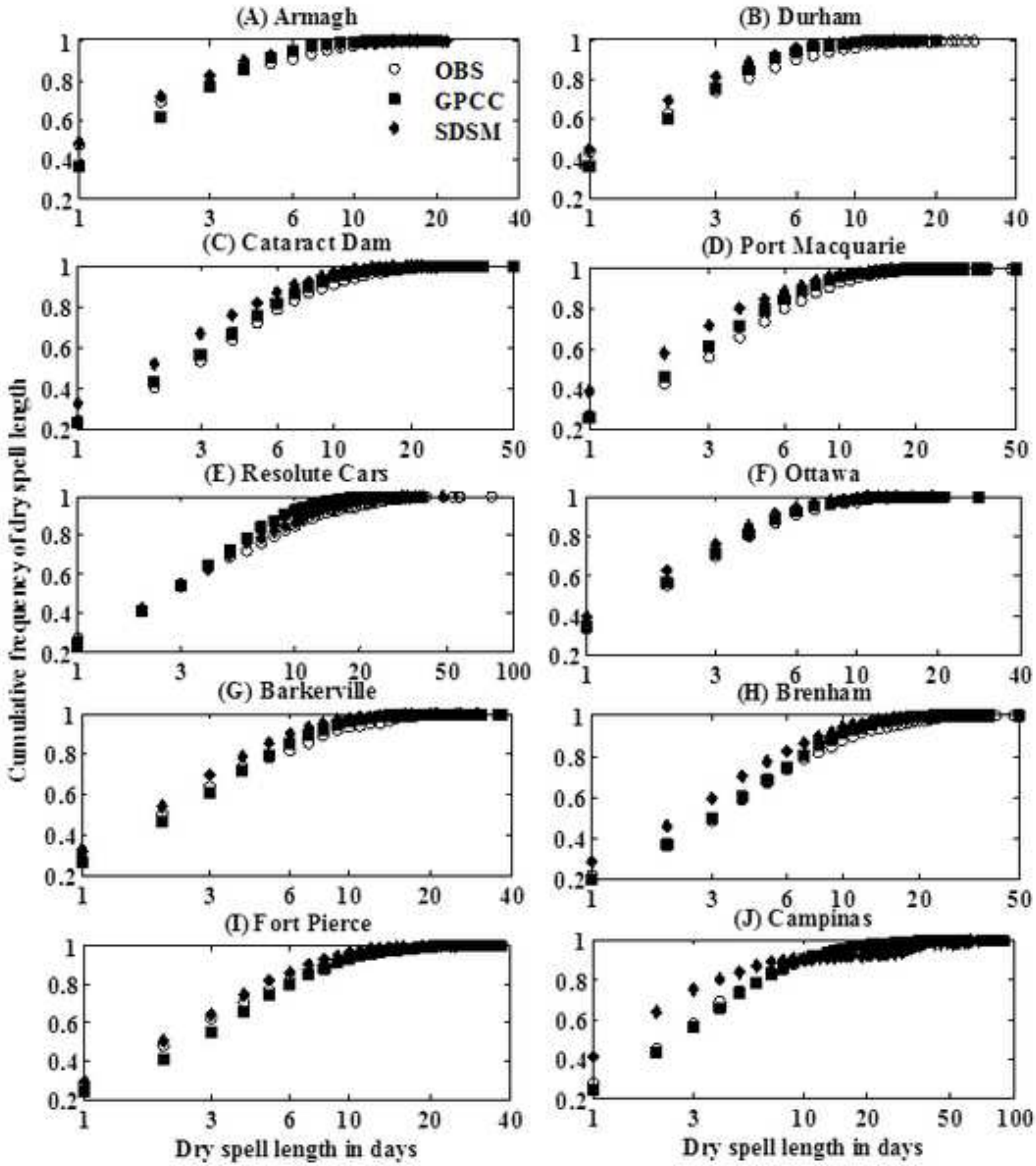

Figure 2 caption

Figure 2. Observed (OBS), GPCC- and SDSM-downscaled cumulative frequencies of dry spells for 10 stations.

Figure 3
Click here to download Figure: Figure3.tif

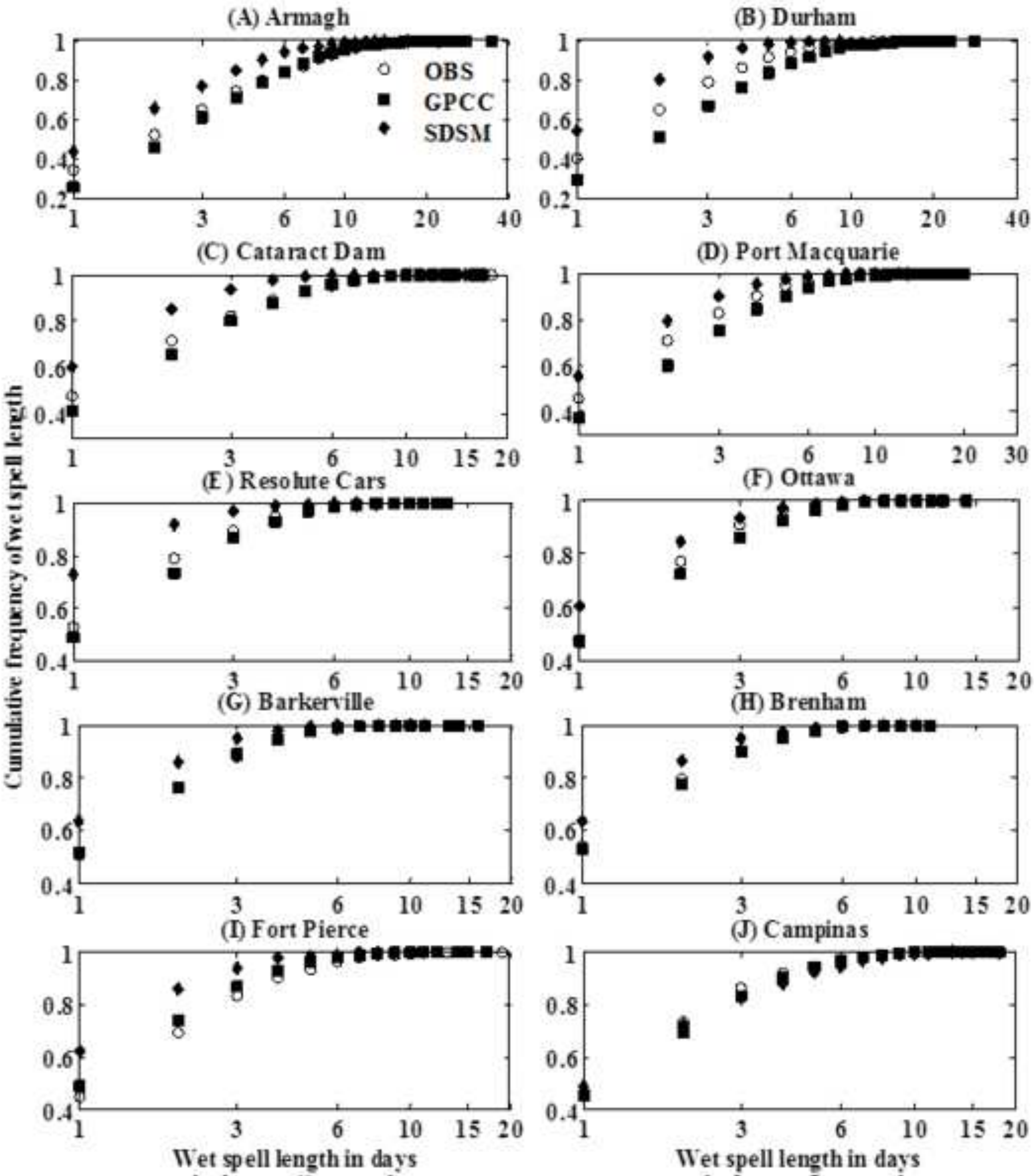

Figure 3 caption

Figure 3. Observed (OBS), GPCC- and SDSM-downscaled cumulative frequencies of wet spells for 10 stations.

Table 1

Model	Issue	Main GPCC Advantage	Main SDSM Advantage
GPCC	Downscales directly from surface climate variables, e.g. precipitation	• Less data intensive and time consuming to downscale from surface variables than screening multiple large-scale predictors	• Large-scale atmospheric predictor variables better simulated by GCMs than surface variables
SDSM	Downscales from large-scale atmospheric climate variables, e.g. geopotential heights		
GPCC	Temporally downscales monthly projections to daily projections using a weather generator	• Monthly projections more reliable than daily projections and are more readily available from many GCMs and emission scenarios	• No temporal downscaling step means no issue with impact models that require information on daily climate characteristics
SDSM	Downscales at a daily resolution = daily projections with no temporal downscaling step		

Table 1. Key advantages and disadvantages of the GPCC and SDSM approaches.

Table 2

Station & Location	Lat. (°E) & Long. (°N)	Timespan	Calibration Period Timespan	Calibration Period MAP (mm)	Validation Period Timespan	Validation Period MAP (mm)	Change (%)
1 Resolute Cars, Canada	-94.98, 74.72	1948-2009	1948-84, 2005-09	135.4	1985-2004	177.6	31.1
2 Barkerville, Canada	-121.50, 53.10	1948-2009	1948-76, 1996-2009	506.0	1977-95	460.4	-9.0
3 Durham, England, UK	-1.57, 54.77	1948-1998	1965-98	651.7	1948-64	627.9	-3.7
4 Armagh, N. Ireland, UK	-6.65, 54.35	1948-2009	1948-54, 1975-2009	793.7	1955-74	845.3	6.4
5 Ottawa, Canada	-75.7, 45.41	1948-2008	1948-51, 1972-2008	920.9	1952-71	805.6	-12.5
6 Brenham, Texas, USA	-96.40, 30.16	1948-2008	1948-88	1017.5	1989-2008	1190.0	17.0
7 Cataract Dam, Australia	150.8, -34.27	1948-2006	1968-2006	1078.4	1948-67	1340.4	24.3
8 Campinas, Brazil	-47.0, -22.83	1948-2010	1948-81, 2002-10	1339.0	1982-2001	1451.3	8.4
9 Fort Pierce, Florida, USA	-80.35, 27.46	1948-2008	1948-70, 1991-2008	1424.7	1971-90	1248.2	-12.4
10 Port Macquarie, Australia	152.86, 31.44	1948-2008	1948-88	1594.4	1989-2008	1382.9	-13.3

Table 2. Details of climate stations, record lengths and precipitation statistics for the calibration and validation period. Numbers next to the station correspond to the numbers shown in Figure 1.

Table 3

Data/Method	GPCC	SDSM
Input data	1. Monthly station precipitation 2. NCEP monthly precipitation	3. Daily station precipitation 4. NCEP daily large-scale predictors
Spatial Downscaling	Quantile mapping between 1 and 2 for calibration period	Transfer functions developed between 3 and 4 for calibration period on monthly basis
Temporal Downscaling	Linear relationships between daily station data and monthly downscaled data used to adjust transition probabilities of precipitation occurrence as input to CLIGEN weather generator	Transfer functions forced with NCEP large-scale predictors used in calibration for validation period as input to SDSM weather generator
Validation	100 year CLIGEN series of daily data developed for validation period and compared with observed daily station data for validation period	20 year series of daily data developed for validation period and compared with observed daily station data for validation period

Table 3. General Overview of the modelling procedure between the two models used in this study.

Table 4

		ARM	DUR	CAT	POR	RES	OTA	BAR	BRE	FOR	CAM
Over	DJF	0.11	0.14	0.38	0.32	0.38	0.47	0.24	0.41	0.45	0.18
	MAM	0.12	0.22	0.48	0.40	0.50	0.39	0.25	0.36	0.35	0.23
	JJA	0.16	0.17	0.50	0.46	0.40	0.23	0.27	0.34	0.26	0.30
	SON	0.14	0.22	0.39	0.33	0.37	0.43	0.18	0.39	0.34	0.19
IDW	DJF	0.31	0.29	0.31	0.32	0.41	0.43	0.24	0.37	0.41	0.21
	MAM	0.25	0.31	0.40	0.33	0.45	0.37	0.24	0.27	0.32	0.25
	JJA	0.26	0.23	0.43	0.36	0.42	0.26	0.26	0.31	0.25	0.30
	SON	0.24	0.31	0.32	0.29	0.34	0.40	0.15	0.36	0.34	0.23

Table 4. Site-specific correlation coefficient (Pearson's r) between daily station precipitation and daily generated precipitation series for the validation period when models are calibrated with the optimum five predictors for each station and season. Over: Overlying grid box; IDW: average of four nearest grid boxes. DJF: Winter; MAM: Spring; JJA: Summer; SON: Autumn. Grey shaded boxes indicate which grid box was selected for subsequent downscaling.

Table 5

ARM	DUR	CAT	POR	RES	OTA	BAR	BRE	FOR	CAM
g1000 r1000 u500 v1000	r500 u500 v850 z850	g1000 s500 u1000 u850	r1000 u850 z850	s850 v1000 z500	s500 u1000 z850	u500 v500 z500 z850	r1000 u500 z1000 z500	u1000 z500 z850	g850 z500

Table 5. Selected predictors for downscaling at each station. G: geopotential height; r: relative humidity; s: specific humidity; u: zonal velocity; v: meridional velocity; z: vorticity; Numbers represent atmospheric pressure level (hPa).

Table 6

	Source	Mean	Stdev	Skew	Kurt	Q25	Q50	Q75	Q90	Q95	Q99	MAX	Sum
Station	OBS	Absolute values											
	GPCC/ SDSM	Relative Error (RE) = Observed – Simulated / Observed											
Mean ARE	OBS v GPCC/ SDSM	Mean of the Absolute Relative Error (ARE). This is the total relative error and does not consider direction of bias											
Mean RE		Mean RE calculated across all stations											

Table 6. Outline of the statistical analysis used to validate GPCC and SDSM for the validation period of each station. This analysis is conducted for MDP, MMP, MAP and AMDP.

Table 7

	Source	Mean	Stdev	Skew	Kurt	Q25	Q50	Q75	Q90	Q95	Q99	MAX	MWD
Armagh	OBS	4.0	5.1	4.0	34.3	0.9	2.4	5.1	9.4	13.5	24.1	78.3	208.9
Durham	OBS	3.7	5.0	3.4	21.1	0.8	2.0	4.7	8.9	12.9	24.5	55.6	167.4
Cataract Dam	OBS	11.1	21.1	4.7	34.0	1.3	3.6	11.2	28.4	49.3	116.9	266.7	121.3
Port Macquarie	OBS	11.1	18.6	4.3	30.9	1.4	4.4	12.4	28.0	45.7	89.1	220.0	124.8
Resolute Cars	OBS	1.9	2.4	3.8	29.7	0.6	1.0	2.2	4.6	6.8	12.2	35.0	91.7
Ottawa	OBS	5.8	7.2	2.8	14.5	1.0	3.0	7.6	14.0	19.3	35.6	71.1	139.5
Barkerville	OBS	3.7	4.2	2.7	13.5	1.0	2.3	5.0	8.8	12.0	20.4	38.8	122.5
Brenham	OBS	12.8	19.7	4.4	38.0	1.8	5.1	16.0	33.8	48.2	90.2	263.7	92.7
Fort Pierce	OBS	9.3	14.6	3.9	30.7	1.0	3.6	11.4	24.9	38.8	65.5	216.7	133.6
Campinas	OBS	13.0	15.3	2.4	12.3	2.3	7.6	18.2	33.0	44.0	66.6	144.7	111.7
Armagh	GPCC	-19.9	0.3	-10.9	-31.4	-66.7	-54.2	-21.6	-4.3	-2.2	-0.5	0.3	2.9
Durham	GPCC	-7.8	22.5	21.9	42.8	-62.5	-50.0	-19.1	4.5	11.6	22.6	69.8	20.6
Cataract Dam	GPCC	-21.3	3.5	26.3	73.1	-76.9	-77.8	-43.8	-11.2	-9.5	-10.1	65.5	11.1
Port Macquarie	GPCC	10.6	38.7	16.3	41.9	-78.6	-22.7	-5.6	24.0	23.3	40.9	126.9	22.2
Resolute Cars	GPCC	-26.0	-12.0	19.3	29.7	-50.0	-40.0	-31.8	-26.1	-23.3	-15.8	12.6	26.7
Ottawa	GPCC	-9.6	7.9	48.8	200.9	-70.0	-16.7	-13.2	-3.6	1.6	2.6	171.0	9.1
Barkerville	GPCC	-2.1	18.2	20.8	40.1	-70.0	-4.3	-16.0	4.5	10.8	18.8	70.9	1.8
Brenham	GPCC	-4.7	-12.2	-16.4	-32.0	-10.0	28.0	-3.1	-9.4	-10.1	-5.2	-4.5	9.7
Fort Pierce	GPCC	0.0	8.8	5.1	1.7	-70.5	6.9	-1.1	3.5	1.2	13.8	20.0	-11.6
Campinas	GPCC	-4.7	-29.6	109.8	302.5	234.8	42.1	-28.6	-37.6	-34.3	-15.7	27.2	3.4
Mean ARE	**GPCC**	**10.7**	**15.4**	**29.6**	**79.6**	**79.0**	**34.3**	**18.4**	**12.9**	**12.8**	**14.6**	**56.9**	**11.9**
Mean RE	**GPCC**	**-8.5**	**4.6**	**24.1**	**66.9**	**-32.0**	**-18.9**	**-18.4**	**-5.6**	**-3.1**	**5.1**	**56.0**	**9.6**
Armagh	SDSM	-0.3	-22.2	-39.0	-62.0	46.7	16.5	5.6	-4.7	-12.8	-22.4	-25.5	-8.3
Durham	SDSM	5.7	-17.7	-21.5	-20.7	55.1	32.1	10.2	-0.3	-8.1	-19.9	44.1	-5.3
Cataract Dam	SDSM	-14.5	-42.3	-29.5	-39.3	58.6	46.6	7.6	-19.4	-34.1	-50.0	-28.1	-5.1
Port Macquarie	SDSM	-0.6	-24.6	-23.9	-29.1	67.9	39.7	14.7	-4.0	-17.6	-25.4	26.1	8.7
Resolute Cars	SDSM	-12.4	-22.8	-4.5	-17.2	-2.5	2.8	-11.7	-22.8	-23.7	-22.5	-9.6	-14.7
Ottawa	SDSM	3.1	-6.1	4.0	22.3	64.9	24.5	1.8	-1.1	-2.0	-9.9	55.0	1.3
Barkerville	SDSM	3.6	-12.6	-7.6	4.5	39.2	21.4	2.4	-4.9	-8.9	-16.0	47.0	-0.9
Brenham	SDSM	-13.9	-24.4	-20.5	-41.0	25.6	16.5	-14.2	-20.6	-19.0	-18.8	-9.9	11.5
Fort Pierce	SDSM	16.1	-0.9	-11.1	-24.1	111.0	63.0	20.2	7.6	-0.9	6.4	18.6	-13.8
Campinas	SDSM	13.8	40.3	24.4	39.0	-25.6	-20.1	4.4	21.2	31.7	54.3	152.3	-1.3
Mean ARE	**SDSM**	**8.4**	**21.4**	**18.6**	**29.9**	**49.7**	**28.3**	**9.3**	**10.7**	**15.9**	**24.5**	**41.6**	**7.1**
Mean RE	**SDSM**	**0.1**	**-13.3**	**-12.9**	**-16.8**	**44.1**	**24.3**	**4.1**	**-4.9**	**-9.5**	**-12.4**	**27.0**	**-2.8**

Table 7. Statistics of observed and simulated mean daily precipitation amounts for the validation period for ten climate stations. Light grey shaded cells reflect positive RE (i.e. overestimations) whereas white cells reflect negative RE (i.e. underestimations). MWD: Mean Wet Days.

Table 8

	Source	Mean	Stdev	Skew	Kurt	Q25	Q50	Q75	Q90	Q95	Q99	MAX
Armagh	OBS	70.3	30.4	0.4	2.7	48.0	67.0	90.1	113.2	126.3	146.7	156.8
Durham	OBS	52.1	32.7	1.2	5.0	28.2	44.2	71.0	96.4	110.2	172.1	185.3
Cataract Dam	OBS	111.7	116.4	2.4	9.9	36.9	74.9	153.6	239.9	317.2	629.0	683.2
Port Macquarie	OBS	115.1	89.0	1.0	3.6	48.0	90.3	165.7	246.3	286.6	368.3	446.0
Resolute Cars	OBS	14.4	14.7	1.6	5.6	4.3	9.2	19.8	37.3	44.7	67.1	78.5
Ottawa	OBS	67.1	32.0	0.7	3.1	43.8	61.6	87.9	111.5	124.2	156.0	171.5
Barkerville	OBS	38.1	23.7	1.3	5.4	20.3	34.8	50.7	69.0	80.2	123.3	141.8
Brenham	OBS	99.2	77.4	1.7	7.2	46.5	83.7	127.8	198.9	243.0	402.0	444.8
Fort Pierce	OBS	104.0	74.1	1.2	5.0	47.2	88.4	140.3	204.2	238.8	332.6	444.5
Campinas	OBS	120.9	98.9	0.8	3.1	36.4	96.0	194.9	257.7	313.0	407.6	422.7
Armagh	GPCC	-17.6	-10.6	166.5	72.1	-19.5	-20.1	-19.4	-17.4	-15.9	-1.1	21.1
Durham	GPCC	11.2	0.6	28.4	50.6	19.5	16.0	3.0	4.0	9.5	-8.3	49.8
Cataract Dam	GPCC	-12.5	-23.4	-13.5	-9.5	0.8	-4.9	-15.8	-10.3	-14.8	-30.1	1.6
Port Macquarie	GPCC	35.2	34.2	43.9	59.7	45.1	38.4	28.2	31.3	36.0	52.3	71.1
Resolute Cars	GPCC	-6.2	-28.7	24.9	64.1	57.6	13.7	-14.1	-25.9	-23.2	-25.0	8.9
Ottawa	GPCC	-1.4	12.8	99.1	94.8	-5.8	-3.1	-4.4	-0.6	9.1	20.9	61.2
Barkerville	GPCC	-0.4	3.0	11.9	7.9	2.0	-5.5	-2.1	0.6	8.4	-2.5	19.8
Brenham	GPCC	4.6	-11.0	-13.7	5.5	17.7	7.2	9.9	-3.8	-4.6	-16.7	36.5
Fort Pierce	GPCC	-11.6	-11.2	13.3	20.7	-8.3	-11.1	-10.0	-12.1	-8.8	-6.9	12.6
Campinas	GPCC	-1.5	-8.4	28.2	38.9	36.5	0.4	-13.3	-3.4	-4.8	-8.3	31.3
Mean ARE	**GPCC**	**10.2**	**14.4**	**44.3**	**42.4**	**21.3**	**12.0**	**12.0**	**10.9**	**13.5**	**17.2**	**31.4**
Mean RE	**GPCC**	**0.0**	**-4.3**	**38.9**	**40.5**	**14.6**	**3.1**	**-3.8**	**-3.8**	**-0.9**	**-2.6**	**31.4**
Armagh	SDSM	-8.5	-12.7	57.0	34.6	-4.1	-8.0	-11.5	-11.9	-10.9	-6.3	29.4
Durham	SDSM	0.1	-35.3	-18.2	3.7	32.0	11.6	-10.0	-17.4	-17.0	-34.2	10.8
Cataract Dam	SDSM	-18.8	-46.1	-20.7	0.5	23.9	3.3	-23.2	-28.5	-34.7	-52.1	-7.7
Port Macquarie	SDSM	8.1	-16.1	2.9	25.0	43.8	23.9	-0.6	-8.9	-7.5	-3.6	14.5
Resolute Cars	SDSM	-25.2	-26.7	15.8	35.3	-22.2	-26.4	-27.0	-29.4	-24.9	-28.6	19.1
Ottawa	SDSM	4.4	-1.2	44.8	52.0	8.6	6.0	-0.3	0.9	4.4	4.5	67.7
Barkerville	SDSM	2.7	-21.2	-43.1	-25.8	23.7	7.1	-0.4	-7.3	-8.7	-25.9	21.1
Brenham	SDSM	-4.1	-13.4	-17.0	-18.7	1.2	-3.7	-1.6	-9.0	-4.9	-19.9	8.0
Fort Pierce	SDSM	0.1	-8.1	-5.7	-4.7	14.0	2.2	-0.2	-2.5	-0.3	-6.6	15.9
Campinas	SDSM	12.3	139.1	155.4	131.1	-100.0	-71.5	-33.5	79.5	130.8	142.2	219.2
Mean ARE	**SDSM**	**8.4**	**32.0**	**38.1**	**33.1**	**27.4**	**16.4**	**10.8**	**19.5**	**24.4**	**32.4**	**41.4**
Mean RE	**SDSM**	**-2.9**	**-4.2**	**17.1**	**23.3**	**2.1**	**-5.6**	**-10.8**	**-3.5**	**2.6**	**-3.1**	**39.8**

Table 8. Statistics of observed and simulated mean monthly precipitation amounts for the validation period for ten climate stations. Light grey shaded cells reflect positive RE (i.e. overestimations) whereas white cells reflect negative RE (i.e. underestimations).

	Source	Mean	Stdev	Skew	Kurt	Q25	Q50	Q75	Q90	Q95	Q99	MAX
Armagh	OBS	843.5	107.5	0.5	2.5	759.0	834.9	897.1	1014.2	1053.6	1073.6	1073.6
Durham	OBS	624.9	120.7	0.0	1.6	518.6	622.1	738.7	781.4	793.5	799.2	799.2
Cataract Dam	OBS	1340.4	446.2	0.6	2.4	984.4	1236.0	1682.3	1976.7	2217.8	2293.1	2293.1
Port Macquarie	OBS	1381.5	360.7	0.4	2.5	1161.0	1318.9	1596.5	1933.3	2025.9	2100.6	2100.6
Resolute Cars	OBS	172.3	46.3	1.1	3.4	138.6	158.8	192.9	255.9	277.0	277.0	277.0
Ottawa	OBS	805.6	84.0	0.4	2.9	750.6	806.3	843.3	920.5	966.8	996.8	996.8
Barkerville	OBS	457.7	72.2	0.2	2.3	400.6	469.2	507.1	546.4	581.9	606.0	606.0
Brenham	OBS	1190.0	305.7	0.0	1.8	955.8	1136.5	1460.4	1601.6	1624.6	1640.6	1640.6
Fort Pierce	OBS	1248.2	224.0	0.2	2.4	1101.2	1225.8	1407.4	1545.7	1630.7	1697.0	1697.0
Campinas	OBS	1450.5	243.2	0.8	4.1	1309.5	1425.6	1588.9	1720.2	1950.0	2111.9	2111.9
Armagh	GPCC	-17.6	-23.0	-21.0	7.5	-16.4	-17.5	-17.4	-18.7	-18.3	-17.5	-17.4
Durham	GPCC	11.2	-14.5	-100	91.3	19.8	12.2	4.0	5.0	7.9	22.3	26.3
Cataract Dam	GPCC	-12.5	-33.2	1.9	17.5	-1.3	-12.8	-17.9	-19.1	-23.8	-12.4	-12.2
Port Macquarie	GPCC	35.2	3.9	-9.0	24.8	38.2	39.2	32.3	21.9	23.3	34.8	36.6
Resolute Cars	GPCC	-6.2	-35.0	-72.4	-17.4	1.4	-1.1	-6.5	-20.4	-22.7	-15.4	-11.7
Ottawa	GPCC	-1.4	51.0	37.1	0.2	-5.7	-3.6	2.9	4.8	8.5	12.3	13.2
Barkerville	GPCC	-0.4	-1.0	14.8	32.2	1.9	-3.2	-1.3	0.8	0.4	4.6	7.1
Brenham	GPCC	4.6	-22.7	-100	93.7	15.2	9.5	-4.1	-1.8	0.2	15.3	24.5
Fort Pierce	GPCC	-11.6	-0.5	287.7	38.8	-13.6	-11.2	-14.4	-6.7	-4.5	1.1	4.9
Campinas	GPCC	-1.5	-24.6	-69.9	-32.3	-2.5	1.4	-3.2	-2.8	-10.6	-9.5	-7.1
Mean ARE	**GPCC**	**10.2**	**20.9**	**71.4**	**35.6**	**11.6**	**11.2**	**10.4**	**10.2**	**12.0**	**14.5**	**16.1**
Mean RE	**GPCC**	**0.0**	**-10.0**	**-99.9**	**25.6**	**3.7**	**1.3**	**-2.6**	**-3.7**	**-4.0**	**3.6**	**6.4**
Armagh	SDSM	-8.5	-14.0	-71.9	-1.4	-6.8	-8.5	-6.5	-11.4	-12.5	-8.6	-4.7
Durham	SDSM	0.1	-37.5	-100.0	72.1	10.4	-0.3	-8.9	-7.2	-3.7	0.4	5.9
Cataract Dam	SDSM	-18.8	-42.5	109.0	101.3	-7.9	-15.8	-28.4	-28.9	-26.3	-15.2	-5.7
Port Macquarie	SDSM	8.1	-25.0	-43.4	7.6	12.3	12.2	4.8	-3.8	-3.7	1.4	10.9
Resolute Cars	SDSM	-25.2	-42.2	-30.2	7.5	-20.6	-21.4	-26.7	-35.0	-35.2	-25.7	-18.3
Ottawa	SDSM	4.4	36.2	-23.4	-5.9	1.2	2.7	8.5	8.7	7.3	11.6	16.3
Barkerville	SDSM	2.7	-24.2	-171.7	26.2	8.2	1.3	-0.2	-1.7	-3.5	-2.6	0.6
Brenham	SDSM	-4.1	-0.8	-100.0	98.6	-2.9	-2.3	-10.5	-4.5	7.6	22.1	37.9
Fort Pierce	SDSM	0.1	-1.8	89.3	44.3	-0.4	1.0	-1.5	-1.9	-2.5	9.5	21.2
Campinas	SDSM	12.3	-7.3	-83.2	-31.7	12.9	14.8	11.6	11.9	4.5	1.6	7.6
Mean ARE	**SDSM**	**8.4**	**23.2**	**82.2**	**39.7**	**8.4**	**8.0**	**10.8**	**11.5**	**10.7**	**9.9**	**12.9**
Mean RE	**SDSM**	**-2.9**	**-15.9**	**-42.5**	**31.9**	**0.7**	**-1.6**	**-5.8**	**-7.4**	**-6.8**	**-0.6**	**7.2**

Table 9. Statistics of observed and simulated mean annual precipitation amounts for the validation period for ten climate stations. Light grey shaded cells reflect positive RE (i.e. overestimations) whereas white cells reflect negative RE (i.e. underestimations).

Table 10

	Source	Mean	Stdev	Skew	Kurt	Q25	Q50	Q75	Q90	Q95	Q99	MAX
Armagh	OBS	37.7	16.2	1.4	4.4	25.6	33.3	44.6	62.8	77.8	78.3	78.3
Durham	OBS	31.8	10.4	0.7	2.7	23.9	27.7	39.6	45.9	52.4	55.6	55.6
Cataract Dam	OBS	138.6	53.4	0.5	3.0	110.1	128.3	171.7	206.0	238.8	266.7	266.7
Port Macquarie	OBS	113.3	47.9	0.8	3.0	80.3	105.7	141.7	188.6	216.1	220.0	220.0
Resolute Cars	OBS	13.7	5.7	2.7	11.3	11.0	13.0	15.3	16.0	25.5	35.0	35.0
Ottawa	OBS	44.6	10.3	0.7	3.6	37.5	44.0	48.6	58.0	66.1	71.1	71.1
Barkerville	OBS	24.5	7.3	0.0	2.0	17.4	25.8	29.6	32.6	36.2	38.8	38.8
Brenham	OBS	112.5	56.4	1.9	5.6	80.3	93.6	120.9	204.0	262.0	263.7	263.7
Fort Pierce	OBS	86.2	41.3	1.9	6.3	61.0	73.0	97.7	142.1	188.1	216.7	216.7
Campinas	OBS	86.1	25.1	1.0	3.2	69.8	78.7	103.8	127.0	141.5	144.7	144.7
Armagh	GPCC	-6.5	-24.7	-16.6	-0.2	4.5	-3.2	-6.1	-13.7	-24.1	-1.9	0.3
Durham	GPCC	39.4	38.2	41.1	64.0	43.2	53.8	30.4	33.1	33.8	65.0	69.8
Cataract Dam	GPCC	12.1	32.7	176.7	92.4	-3.6	6.6	7.8	18.2	19.9	52.7	65.5
Port Macquarie	GPCC	60.4	54.7	100.3	120.1	58.6	55.4	58.4	43.5	42.6	117.0	126.9
Resolute Cars	GPCC	-3.0	0.8	-30.1	-28.9	-14.1	-6.2	-0.3	23.4	-8.8	4.3	12.6
Ottawa	GPCC	11.7	113.8	361.8	448.0	-1.6	0.5	18.0	22.3	36.2	113.4	171.0
Barkerville	GPCC	19.0	27.8	3420.8	180.8	30.2	5.0	13.5	27.0	25.4	65.6	70.9
Brenham	GPCC	-10.3	-31.0	-11.6	7.8	-6.1	-2.1	-6.1	-25.7	-29.0	-8.7	-4.5
Fort Pierce	GPCC	11.6	-0.9	-14.1	-8.4	15.7	16.9	11.5	8.1	0.2	12.8	20.0
Campinas	GPCC	-9.7	37.8	14.2	28.5	-27.2	-12.3	-8.1	0.0	8.1	26.4	27.2
Mean ARE	**GPCC**	**18.4**	**36.2**	**418.7**	**97.9**	**20.5**	**16.2**	**16.0**	**21.5**	**22.8**	**46.8**	**56.9**
Mean RE	**GPCC**	**12.5**	**24.9**	**404.3**	**90.4**	**10.0**	**11.4**	**11.9**	**13.6**	**10.4**	**44.7**	**56.0**
Armagh	SDSM	-33.3	-57.2	-5.2	30.1	-21.0	-28.8	-36.0	-44.8	-50.0	-39.8	-25.5
Durham	SDSM	-21.4	-24.4	210.7	424.6	-17.7	-13.6	-28.6	-25.9	-26.0	-8.9	44.1
Cataract Dam	SDSM	-47.7	-53.4	135.6	63.1	-50.9	-48.2	-50.7	-48.4	-49.6	-43.9	-28.1
Port Macquarie	SDSM	-23.4	-36.6	109.4	155.7	-17.9	-24.7	-29.3	-32.7	-32.3	-13.8	26.1
Resolute Cars	SDSM	-23.4	-31.4	-43.4	-42.0	-29.8	-24.8	-21.3	-4.7	-27.9	-31.6	-9.6
Ottawa	SDSM	-7.8	25.5	144.6	115.2	-14.0	-12.9	-4.0	-2.1	-0.3	26.5	55.0
Barkerville	SDSM	-13.9	-11.7	4419.4	292.2	-4.8	-22.4	-19.8	-11.5	-8.7	16.9	47.0
Brenham	SDSM	-22.0	-44.4	-39.0	-14.3	-18.0	-12.8	-14.4	-37.3	-42.5	-30.1	-9.9
Fort Pierce	SDSM	0.6	-22.9	-20.1	11.9	5.1	8.0	5.3	-9.1	-23.6	-0.7	18.6
Campinas	SDSM	40.6	48.5	82.3	209.2	39.3	45.1	31.6	29.2	35.2	68.0	152.3
Mean ARE	**SDSM**	**23.4**	**35.6**	**521.0**	**135.8**	**21.9**	**24.1**	**24.1**	**24.6**	**29.6**	**28.0**	**41.6**
Mean RE	**SDSM**	**-15.2**	**-20.8**	**499.4**	**124.6**	**-13.0**	**-13.5**	**-16.7**	**-18.7**	**-22.6**	**-5.7**	**27.0**

Table 10. Statistics of observed and simulated annual maximum daily precipitation amounts for the validation period for ten climate stations. Light grey shaded cells reflect positive RE (i.e. overestimations) whereas white cells reflect negative RE (i.e. underestimations).

		ARM	DUR	CAT	POR	RES	OTA	BAR	BRE	FOR	CAM
	OBS	19	41	38	48	81	18	30	48	29	84
Dry	GPCC	20	20	50	50	37	28	37	50	37	88
	SDSM	28	26	40	43	59	32	32	55	44	79
	OBS	23	16	18	13	10	9	11	10	19	15
Wet	GPCC	35	28	17	20	13	14	16	11	17	18
	SDSM	23	15	15	18	12	14	12	16	14	24

Table 11. The longest dry and wet spells (days) extracted from observed, GPCC- and SDSM-downscaled daily precipitation series for the validation period for all ten stations. Dark grey = overestimations; light grey = underestimations; white = no change. Station acronyms represent the ten stations in order in Tables 2-6.